

Neurala korrelat till fyllda pauser: En fMRI-studie av disfluensperception

Robert Eklund

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-155669>

N.B.: When citing this work, cite the original publication.

Eklund, R., (2019), Neurala korrelat till fyllda pauser: En fMRI-studie av disfluensperception, *Röstläget*, (Februari), 13-17.

Original publication available at: <https://rostforum.se/>

Copyright: The Author

URL: <https://rostforum.se/>

This article is an extended and more detailed version of the following article:
Neurala korrelat till fyllda pauser. En fMRI-studie av disfluensperception



Neurala korrelat till fyllda pauser. En fMRI-studie av disfluensperception

Robert Eklund

Biträdande Professor i Språk, Kultur och Kommunikation

Institutionen för Kultur och Kommunikation, Linköpings Universitet

&

Docent i Datorlingvistik

Institutionen för Datorlingvistik, Linköpings Universitet

<http://roberteklund.info>

Not:

Denna artikel är en lång och mer utförlig version av följande artikel:

Neurala korrelat till fyllda pauser. En fMRI-studie av disfluensperception

Röstläget, februari 2019 sid. 13–16 samt sidan 19 (figurer).

Artikeln sammanfattar det plenarföredrag som hölls på *Röstforum 2012*, Linköpings Universitet.

Av främst publiceringsstrategiska skäl blev artikeln försenad.

Introduktion

Mänskligt, spontant producerat (inte uppläst eller på annat sätt planerat eller inrepeterat) talspråk kännetecknas av att inte vara helt ”flytande”. Jag sätter ordet inom citationstecken eftersom det råder delade meningar om huruvida ”oflyt” i själva verket underlättar såväl talproduktion som talperception. Den vanligaste termen för detta är *disfluenser*, men även denna term är inte helt etablerad. En annan sak att hålla i minne är att den alternativa stavningen *dysfluenser* förekommer, speciellt engelskspråkig litteratur. Eftersom de grekiska prefixen *dis-* respektive *dys-* betyder olika saker, där det förra helt enkelt är en negation och det senare anger att patologiskt tillstånd, borde det korrekta bruket vara att reservera *disfluenser* för icke-patologiskt oflyt och *disfluenser* för patologiskt oflyt, men termerna förekommer ofta synonymt, säkerligen (åtminstone delvis) på grund av bristen på en fonologisk distinktion mellan /i/ och /y/ i engelska. Detta kan till och med leda till självmotsägande begrepp som ”normal dysfluency” (Ratner & Sih, 1987:278; Dale, 1977:312; Shapiro & DeCicco, 1982:109; Westby, 1974:133). Se Wingate (1984) samt Eklund (2004: 157–160) för en vidare diskussion av ytterligare alternativa termer och begrepp, samt terminologisk förbistring.

Disfluenser har studerats i över ett sekel, och en introduktion följer nedan. Denna artikel redovisar resultaten från en unik fMRI-studie av den mest ”speciella” (förklaring följer) av de olika disfluentstyperna, det som ofta(st) benämns ”fyllda pauser”, dvs vokaliseringar (eller, vill vissa göra gällande, ”ord”) som (i svenska) ”eh” eller ”öh”. Notera att även denna term inte är etablerad.

Icke-patologiska disfluenser

Ungefär 6 % av spontant tal utgörs av något slags disfluens, dvs vi ”stakar oss” på ett eller annat sätt på (i snitt) ungefär vart 20:e ord. Denna siffra har visat sig vara tämligen stabil mellan de språk som varit huvudsakligt fokus för disfluensstudier (se t ex [Fox Tree, 1995](#); [Oviatt, 1995](#); [Brennan & Schober, 2001](#); [Bortfeld et al., 2001](#); [Eklund, 2004](#)). Formella studier av disfluenser påbörjades på 1930-talet av Wendell Johnson och kollegor ([Johnson and Associates, 1948, 1955, 1959, 1961](#)), med som huvudsakligt syfte att reda ut orsaker till och prediktion av stamning,, men studerades de kommande decennierna, med andra syften, inom andra discipliner, och med banbrytande studier främst inom psykoterapi ([Mahl, 1956, 1958, 1987a, 1987b](#); [Kasl & Mahl, 1958, 1965, 1987](#); [Goldman-Eisler, 1954a, 1954b, 1955, 1957](#)) och allmän lingvistik ([Hegedüs, 1953](#); [Goldman-Eisler, 1954, 1955, 1957, 1972](#); [Maclay & Osgood, 1959](#)). Disfluenser har även studerats inom en stor mängd andra discipliner som neurovetenskap, psykologi, sociolingvistik, filosofi, medvetandefilosofi, teckenspråk, språkteknologi, paralingvistik, tvåspråkighetsforskning, biologi, etc. För en omfattande redogörelse för och sammanfattning av forskningen upp till år 2004 hänvisas till [Eklund \(2004:51–171\)](#).

Rörande exakt vilka språkliga fenomen som räknas som disfluenser så varierar även detta inom litteraturen, kanske främst rörande huruvida företeelser vilka i andra sammanhang går under namnet *diskursmarkörer*, som *va*, *asså*, *ba'*, *liksom* etc skall inkluderas under paraplytermen disfluenser. Eftersom samtliga sådana fenomen har ”riktiga” ord som bas – de kan t ex slås upp i ordböcker – har jag – och jag är inte ensam om det – valt bort sådana företeelser, och räknar enbart följande fenomen, med ungefärliga proportioner av det totala antalet disfluenser angivna i procent (siffrorna är tagna från [Eklund \(2004\)](#): *ofyllda pauser* (tystnader, UP), >50 %; *fyllda pauser*, ~25 %; *prolongationer* (segmentförlängningar), ~6 %; *trunkeringar* (avhuggna ord), ~6 %; *feluttal*, ~1 %; och ”explicit redigering”, som ”hoppсан”, ”nej, det var fel” (etc), <1 %. Notera att denna sista kategori utgör ett undantag till regeln om ”inga riktiga ord”, men jag inkluderade denna eftersom det på ett tidigt stadium i den språkteknologiska utvecklingen (datorer som förstår talspråk, så kallad *Automatic Speech Recognition* (ASR) – som jag var med om att utvecklar på Telia Research under andra halvan av 1990-talet – föreslogs att datorer skulle kunna använda sig av (se t ex [Hindle, 1983](#)) för att lättare hantera spontant producerat talspråk.

Med start år 1999 har det arrangerats (någorlunda regelbundna) internationella workshops med icke-patologiska disfluenser som fokus, med namnet *Disfluency in Spontaneous Speech (DiSS)*, med möten i Berkeley 1999, Edinburgh 2001, Göteborg 2003, Aix-en-Provence 2005, Tokyo 2010, Stockholm 2013, Edinburgh 2015 och Stockholm 2017 (och med nästa workshop planerad till september 2019 i Budapest). Proceedings från samtliga dessa workshops kan laddas ned från min hemsida (länk i slutet av denna artikel) – där DiSS 2017:s hemsida ligger speglad.

Fyllda pauser

Som vi såg ovan så utgör fyllda pauser (FP i följande text) den vanligaste formen av vokaliserade disfluenser. Ofyllda pauser undviks tämligen ofta i analyser eftersom det är svårt att dra en gräns för när en tystnad i spontalt talspråk, ”spontal”, faktiskt är en disfluens. Dessutom används ofyllda pauser ofta på ett medvetet, retoriskt sätt, där durationen i sig kan utgöra ett medvetet funktionellt grepp. Mitt favoritexempel ges av ståuppkomikern Jerry Seinfeld, som konstaterade att ”mitt-tystnaden” i frågekonstellation ”Could I ask you a favour ... [tystnad i N sekunder]... could you X ”, där X är den tjänst man ber om och N är tystnadens duration. Ju högre N , desto större X . Lingvistiskt mycket insiktsfullt, anser jag.

FP utgör ungefär en fjärdedel av samtliga disfluenser, vilket innebär att ungefär 3,5 % av alla ”ord” vi yttrar är ett ”eh”. En sammanställning av FP-frekvens, på ordnivå, presenteras i Tabell 1.

Tabell 1: FP-frekvens i olika källor, angivet om procent av det totala antalet ord. Tabellen är en sammanställning av tabeller i [Eklund \(2010\)](#). Legend: H–M: Människa–Maskin-konversation. H–H: människa–människa-konversation. ”Push-to-talk”: talarna pratade inte förrän de var redo, och tryckte på en knapp alldeles innan de började tala, och hade alltså antagligen förberett sitt tal i förväg. WOZ: Wizard-of-OZ: en datainsamlingsmetod där man låter försökpersonerna tro att de pratar med ett automatiskt system (en dator), vilket i själva verket utgörs av en eller flera människor som ”bakom draperiet” låtsas vara systemet; för en utförlig beskrivning av denna metod, se [Eklund \(2004:179–180\)](#); [Allwood och Haglund \(1992\)](#).

Källa	FP/ordnivå	Korpustyp
Maclay & Osgood (1959)	3.9 %	Konferensdata
Lalljee & Cook (1969)	2.8 %	”High pressure” (hög stressnivå)
	3.4 %	”Low pressure” (låg stressnivå)
Lalljee & Cook (1973)	1.9 %	Kvinnor
	4.3 %	Män
Lutz & Mallard (1986)	3.6 %	Konversation
Eklund & Shriberg (1998)	3.4 %	H–M WOZ
	2.6 %	H–H
	0.3 %	H–M ”push-to-talk”
	3.0 %	H–H
Bortfeld et al. (2001)	2.6 %	Konversation
Eklund (2004)	3.0 %	H–M WOZ
	4.1 %	H–M WOZ
	2.2 %	H–H
	4.4 %	H–M
Moniz, Mata & Viana (2007)	2.3 %	Presentationer
Eklund (2010)	7,5 %	Autentiska kundsamtal till Telia 90 200 med riktiga ärenden. Talarna var inte medvetna om att de spelades in.

Som visas i Tabell 1 så är FP-frekvensen tämligen stabil i olika studier. De två resultat som sticker ut är den låga siffran 0,3 % i [Eklund och Shriberg \(1998\)](#), vilken kan förklaras med att just den korpusen utgjordes av ”push-to-talk”-talspråk, dvs försökspersonerna pratade inte förrän de ”var redo”, och först då tryckte på en knapp för att tala, samt den höga siffran 7,5 % ([Eklund, 2010](#)), som intressant nog är den enda av de rapporterade korpusarna som baseras på taldata där försökspersoner dels inte utförde artificiella uppgifter (typ experiment) utan autentiska ärenden, dels (till skillnad från de andra korpusarna) inte visste att de spelades in. Således uppvisade datamängden i Eklund (2010) en ekologisk validitet som både var omatchad, och troligtvis kommer att förbli omatchad. (Notera att datainsamlingen var fullkomligt laglig.) Eftersom det har visats att FP-produktion ligger inom mänsklig kontroll (se t ex [Siegel, Lenske & Broen, 1999](#)) så är det inte otänkbart att försökspersonerna i de övriga studierna i viss mån undertryckte FP-produktionen. Den rapporterade skillnaden mellan kvinnor och män som [Lalljee och Cook \(1973\)](#) rapporterar – och som upprepats i andra studier av disfluensproduktion i amerikansk engelska – har inte noterats i studier av svenska ([Bell, Eklund & Gustafson, 2000](#)).

Men det riktiga intressant med FP är att det tidigt visade sig att FP uppvisar såväl en egen distribution som beteende jämfört med alla andra disfluentstyper, vilket fick [Mahl \(1858\)](#) att skilja på ”*ah* disfluencies” och ”*non-ah* disfluencies”, en distinktion som stått sig i över ett halvsekel (se t ex [Christenfeld & Creager, 1996](#)) och fortfarande är valid. För en utförligare redogörelse för skillnader mellan FP och andra disfluenser hänvisas till [Eklund \(2004\)](#).

Under de decennier som har förflutit sedan Mahl först observerade skillnaden mellan FP och alla andra disfluenser har ett antal olika hypoteser rörande orsaken till och/eller funktionen för FP föreslagits. Jag presenterar nedan en sammanställning av några av de vanligaste hypoteserna. Notera att denna listning är långt ifrån fullständig samt att namnen på hypoteserna inte är ”officiella” eller etablerade utan är mina egna förslag, som presenterade i [Eklund och Wirén \(2010\)](#).

Floor-holding hypothesis. [Maclay & Osgood, \(1959\)](#) var antagligen först ut med att föreslå att FP kan användas för att förhindra att man blir avbruten av en annan talare, dvs man ”behåller golvet” i en konversation, även om man vid vissa tillfällen inte har nästa ord omedelbart färdigt för leverans. Det är helt enkelt svårare att avbryta någon som säger ”eh” än någon som blir helt tyst. Denna hypotes föreslogs även av [Livant \(1963\)](#).

Help-me-out hypothesis. En alternativ hypotes som går ut på motsatsen är att ”eh” utgör ett slags kommunikativ ”flagga” att talaren behöver hjälp av interlokutören/rna. Detta föreslogs av [Clark & Wilkes-Gibbs \(1986\)](#). En mildare form av denna hypotes är att FP helt enkelt signalerar att talaren för tillfället har problem med sin talproduktion, som föreslaget av [Clark och Fox Tree \(2002\)](#). Så, när en talare söker för ett ord – eller helt enkelt för ett sätt att fortsätta prata – så är ”eh” ett tecken på att ”hjälp önskas”.

Self-monitoring/error detection hypothesis. Levelt (1989) föreslog att FP är ett tecken på intern feldetektion, en tråd som följdes upp av Christenfeld och Creager (1996) som var av meningen att allting som resulterade i problematisk talproduktion kunde resultera i emitterade FP. Ett exempel är det ”choking under pressure” Baumeister (1984) föreslagit.

Many-options hypothesis. Lounsbury (1954:99) föreslog att FP “correspond to the points of highest statistical uncertainty in the sequencing of units in any given order”, dvs att man vid början av yttrande – där entropin är som högst – innan talaren ännu har ”committat” sig till något, och där talproduktionen således fortfarande är som mest öppen och därmed komplicerad uppvisar en större sannolikhet för att producera ett ”eh”. Detta har över åren blivit bekräftat i ett otal studier, varav ett antal är: Beattie & Barnard (1979) rapporterade att 55.3% av alla FP som producerades av talare i telefonsamtal till nummerupplysningen inledde yttranden. Cook (1971) observerade att FP hade en stark tendens att uppträda som det första, andra eller tredje ordet i (bi)satser. Shriberg (1994) och Eklund och Shriberg (1998) rapporterade att FP förekom i början av yttranden oftare än i någon annan position. Eklund (2004:235) rapporterade att 45,3 % av FP uppträdde yttrandeinitialt (som ett genomsnittsvärde av fyra olika talspråkskorpora). Boomer (1965) observerade att den vanligaste position för tvekljud (~FP) var som det första ordet i en fonemisk sats. Pfeifer och Bickmore (2009:462) rapporterade att 63,5 % ”of all occurrences of *um* were located at the start of a dialogue turn”.

Men den kanske mest iögonenfallande – och min favorit – bekräftelsen av denna hypotes är den studie som Schachter et al (1991) utförde, vilka insåg att olika discipliner upp visar olika grad av inherent entropi (oordning). De besökt sig således för att studera föredrag inom tre olika discipliner, men olika grad av ”oordning”: (1) *naturvetenskap*, där det är lättare att förutsäga vad ”nästa ord” kommer att vara, eftersom det finns en begränsad mängd sätt att avsluta t ex en påbörjad mening om en planetbanas karakteristik, eller utkomsten av en kemisk reaktion; (2) *samhällsvetenskap*, med en ”mellannivå” av förutsägbarhet vad rör en diskussion kring tänkbara svar på hypoteser; (3) *humaniora*, med en mycket hög grad av oförutsägbarhet vad gäller hur en föredragshållare tänker avsluta en påbörjad mening. Som ett exempel ber författarna fundera på hur följande mening avslutas: ”The reason Jackson Pollock put the patch of red in that corner of the canvas was....” (Schachter et al (1991:362; fyra punkter i originalet). Mycket riktigt producerade föredragshållare inom naturvetenskapliga discipliner färst FP, samhällsvetenskapliga områden lite fler, men föredragshållare inom humaniora flest FP. För att utesluta att individuella drag påverkat resultatet (det kan ju tänkas att personer som säger mycket ”eh” dras till humanistiska ämnen?) bad författaren samtliga föredragshållare ge föredrag inom samma, och ”neutrala” områden, och fick då liknande frekvens av FP för samtliga talare.

Attention-getting signal. Lalljee & Cook (1974) skapade ett antal experiment för att testa floor-holding-hypotesen, men fann inget stöd för denna hypotes, vilket förelidde författarna till att förslå att FP i stället, och helt enkelt, utgör ett sätt att få uppmärksamhet. Detta skulle även förklara den höga incidensen av FP i yttrande-initial position ("här är jag, nu är det min tur att prata"). Emellertid, och i mitt tycke mest insiktsfullt, påpekar Lalljee och Cook (1974) även att FP mycket väl kan fylla flera olika funktioner, och att ett experiment som är designat för att undersöka en (datorlingvistik- och logik-jargon för "en och endast en) funktion mycket väl kan missa andra funktioner.

För att summera denna – ofullständiga och förenklande – listning av olika föreslagna funktioner som FP kan tänkas ha är det viktigt att tänka på att FP mycket väl kan ha flera funktioner i talspråk – vilket Lalljee & Cook (1974) påpekade – samt att dessa mycket väl – och här kommer mina egna observationer genom decenniernas disfluensstudier och de åsikter dessa har gett mig – mycket väl både kan vara "sanna utan att lyckas" (vi återkommer till det när jag redovisar resultaten från den fMRI-studie som beskrivs i denna artikel) samt helt enkelt kan vara sanna allihopa, och kanske till och med i vissa fall samtidigt. Men som slutord kan man inte komma ifrån att many-options-hypotesen har rönt en oerhört stor mängd stöd i de studier som utförts. Det verkar helt enkelt vara som Christenfeld (1994:192) konstaterar: *more options did produce more filled pauses*". (Inom parentes har jag i två radiointervjuer, en för SR P1 *Språket*, och en för NRK *Språkteigen*, diskuterat detta utifrån perspektivet radiospråk; länkar för nedladdning står att finna i slutet av denna artikel.)

Fyllda pauser: bra eller dåliga?

Vi har ovan sett att FP har studerats i detalj och över decennier utifrån deras funktion ut ett talproduktionsperspektiv. Varför *produceras* FP? Den uppenbara och följande frågan är vilken effekt FP har på lyssnaren, dvs vad har studier av *perception* av FP uppvisat? (Och vi kommer hela tiden närmare fMRI-resultaten som denna artikel har i titeln.) Jag ska i det följande mycket (!) kortfattat (jag lovar) redogöra för ett antal studier som har rapporterat resultat från perceptionsstudier av disfluent tal (snarare än FP specifikt).

Studier som har visat att disfluenser *kan ha negativa effekter* på lyssnare inkluderar: McCroskey och Mehrley (1969) fann att tal som innehöll disfluenser resulterade i "mindre övertygande tal". Duffy, Hunt Jr. och Giolas (1975) fann att disfluenser negativt påverkade lyssnarens uppfattning av talarens kompetens. Fox Tree (1995) fann att "false starts" (talaren avbryter sig och får börja om från början) försvårade talförståelse, men att repetitioner (upprepade ord) inte hade denna effekt. Lickley och Bard (1996) fann att specifika ord kan vara svårare att förstå i allmänt disfluenta yttrande än i flytande yttranden. (OBS! Vi återkommer till detta nedan, "med en twist".) Christenfeld (1995) rapporterar att även om flytande tal ger bäst intryck så upplevdes det mycket bättre om talaren sade "eh" än om talaren blev tyst (ofyllda pauser). Exakt samma slutsats drogs av Bortfeld et al. (1999). Ergo: vid tvekan, tystna inte utan säg "eh"!

Studier som har visat på positiva effekter (och nu börjar det bli riktigt intressant, som vi ska se senare) inkluderar: [Fox Tree \(2001\)](#) fann att ”uh” fick lyssnare att förstå kommande ord snabbare än i flytande tal – medan den amerikanska varianten ”um” inte hade denna effekt (för ett par studier rörande skillnaden mellan ”uh” och ”um” hänvisas till [Clark & Fox Tree \(2002\)](#) och [O’Connell & Kowal, 2005](#)). OBS! Kom ihåg just denna studie, som är av centralt intresse för den här presenterade fMRI-studien.) [Bailey och Ferreira \(2003\)](#) fann att FP hade en effekt på syntaktisk parsning. Slutligen, och av centralt intresse för den aktuella fMRI-studien så visade [Arnold et al. \(2003\)](#) att FP fick lyssnare mer predisponerade mot perception av inte tidigare nämnda objekt, medan flytande tal fick lyssnare mer predisponerade mot tidigare nämna objekt. Detta bekräftades i en unik EEG-studie av [Corley, MacGregor och Donaldson \(2007\)](#), som visade att den EEG-komponent som kallas N400 (EEG uppvisar en negativ ”spik” efter ca 400 ms), och som är välkänd vid lyssning på tal som innehåller semantiska anomalier, i hög grad undertrycktes om den semantiska anomalin följde en FP.

Fyllda pauser: en sammanfattning

Vi har ovan sett att FP är *vanliga*; att de *skiljer sig från alla övriga slags disfluenser* i en mängd dimensioner (som funktion och effekt); att de *kan tjäna flera syften* (kanske simultant); att de kan ha *såväl ”skadlig” som hjälpsam effekt* på lyssnaren, med mera.

Fyllda pauser: vår frågeställning

En sak som nästan samtliga av ovan nämnda studier – speciellt de kognitivt inriktade – har gemensamt är följande drag:

- De har inte studerat FP *per se*, utan dessas effekt på följande, och ”riktiga” ord i tal (eller liknande).
- De har i hög grad utförts på scriptat, förberett eller på annat sätt ”skapat” tal (med ett notabelt undantag i [Eklund \(2010\)](#)).
- De ”kognitiva” studier som utförts har antingen varit beteendestudier (som [Fox Tree, 2001](#)) eller utförts elektrofysiologiskt/med EEG (som [Corley, MacGregor & Donaldson, 2007](#)).

Vad vi ville studera var i stället:

1. Studera den effekt FP *per se* har på hjärnan, alltså inte den effekt som FP kan ha på följande ord, eller liknande.
2. Vi använde fMRI (se nästa stycke) i stället för beteendestudier eller EEG.
3. I stället för att använda scriptat eller på annat sätt artificiellt producerat tal för våra stimuli använde vi data med högsta ekologiska validitet, som beskrivet i [Eklund \(2010\)](#).

Notera att dessa tre punkter gör vår studie unik – vilket medför såväl för- som nackdelar, som vi ska se i det följande.

Neurokognition

Jag skall inte redogöra för neurokognitionens historia (ett *enormt* område) här, men för att följa mitt plenarföredrag på Röstforum (den 21 april 2012) redogör jag för några av de första milstolparna inom detta område.

Hjärnan har inte alltid associerats med kognitiva processer, och ett belysande exempel är att man i antika Egypten vid mumifiering var noga med att bevara alla organ utom hjärnan, som snabbt avlägsnades och slängdes bort; detta är ett välkänt faktum; men se t ex [Jackowski, Bolliger & Thali. 2008:1481–1482](#)). Att hjärnan hade med såväl kognition som personlighet blev dock alltmer uppenbart, och ett exempel som förekommer flitigt i litteraturen behandlar järnvägsarbetaren Phineas Gage (1823–60) som vid en olycka 1848 fick ett järnspett genom hjärnan, och som mirakulöst överlevde, men med en totalt förändrad personlighet (se [Ward, 2006:331](#)).

Hjärnan är ett hungrigt organ, och trots att den i snitt enbart utgör ca 2 % av kroppsvikten förbrukar den 20 % av den totala mängden syre i kroppen samt 25 % av kroppens glukos, samt svarar som mottagare av 15 % av hjärtats ”output” ([Jain, Langham & Wehrli, 2010](#)). Så den uppenbara frågan är huruvida man kan använda denna ”aptit” för att objektivt mäta kognitiva processer?

En pionjär inom området var Angelo Mosso (1846–1910), som mätte hjärnaktivitet genom att lägga försökspersoner på en perfekt balanserat säng-liknande våg. När hjärnan började processa information strömmade mer blod till hjärnan (för att förse den med syre), vilket gjorde den tyngre, vilket fick sängen (vågen) att tippa över till den sida där huvudet befann sig. Detta finns beskrivet av [William James \(1890:98\)](#):

The subject to be observed lay on a delicately balanced table which could tip downward either at the head or at the foot if the weight of either end were increased. The moment emotional or intellectual activity began in the subject, down went the balance at the head-end, in consequence of the redistribution of blood in his system.

Det är denna ökade blodtillströmning/syreförbrukning som utgör basen för fMRI-studier.

Hjärnavbildning (“neuroimaging”)

Det finns nuförtiden en stor mängd olika sätt att mäta hjärnans aktivitet, och man kan t ex använda hjärnans elektriska aktivitet (EEG), dess magnetiska aktivitet (MEG, TMS) eller, som i den här studien fMRI (functional Magnetic Resonance Imaging), där man använder sig magnetfält och radiovågor för att mäta lokalt ökad aktivitet i olika delar av hjärnan. Observera att man inte mäter områden där ”hjärnan är aktiv” – något som antagligen till viss del ligger grund för den utbredda myten om att ”man använder bara 10 % av hjärnan” (se t ex [Wanjk, 2002](#)). Eller som [Dündar och Günduz \(2016\)](#) sammanfattar det: ”The idea that we use only 10% of our brain’s capacity is the most

common neuromyth /.../ However, science has failed to confirm any such unused region” (Dünder & Günduz:213). För mer (och rolig) diskussion om 10 %-myten hänvisas till [Beyerstein \(1999\)](#), som bl a framför att denna myt antagligen har en del att göra med att vi har tio fingrar.

Det man mäter med fMRI är (marginellt) ökad aktivering i vissa, specifika, områden i hjärnan, och även om den spatials upplösningen är mycket bra så är tidsupplösning på sekundnivå (jämfört med EEG, som är realtid) och signal-brus-förhållandet är som bäst 3 %. För en kortfattad genomgång av olika metoder för att mäta hjärnaktivitet hänvisas till [Ahlsén och Eklund \(2012\)](#). För mer information om fMRI hänvisas till den utsökta webbsidan [fmri4newbies](#) (länk i slutet av denna artikel).

Den aktuella studien

Den studie som beskrivs i denna artikel utfördes som ett postdocarbete på Karolinska Institutet, avdelningen för klinisk neurovetenskap, med Martin Ingvar som huvudhandledare. En stor mängd andra personer var till enorm hjälp och dessa tackas i slutet av denna artikel. Såväl studien, som publiceringen av dess resultat försenade på grund av att huvudförfattaren (jag) avslutande min anställning på Telia Research (när Telia lade ned sin forskningsaktivitet) och därmed inte längre hade legal rätt till de data på vilka studien baserades. Efter att avtal upprättats mellan Karolinska Institutet, TeliaSonera (som det vid det laget hette) och författaren (Robert Eklund) fick vi åter rätt av fullfölja studier, men vi fann vid ett senare läge att tilläggsavtal krävdes för att kunna skicka in artiklar till tidskrifter som inte enbart kräver tillgång till rådata (vilket är legio) utan som dessutom kräver att samtliga rådata publiceras i samband med artikelpublicering. Utöver detta kräver flera tidskrifter att ingenting om studien på minsta sätt finns publicerat innan publicering sker i aktuell tidskrift.

Experimentbeskrivning

Eftersom en detaljerad experimentbeskrivning egentligen kräver en tämligen god kunskap om hur fMRI fungerar hänvisar jag till (t ex) [fmri4newbies](#) (länk nedan) och undviker att här redogöra för mer tekniska detaljer av vårt experiment. För en detaljerad sådan, teknisk, beskrivning hänvisar jag till [Eklund och Ingvar \(2016\)](#), samt ger i slutet av denna artikel en länk för direkt nedladdning av denna artikel (samt den poster som presenterades på Interspeech i San Francisco år 2016).

Nedan följer dock en grundbeskrivning av experimentet (utan tekniska detaljer):

Stimulusdata var ena sidan av HH-dialoger, insamlade i en WOZ-datainsamling. Det rörde sig om ena sidan (”kunden”) av affärsresebeställningar, vilka producerade talet baserat på ordfria instruktionsblad (för att undvika direktkopiering från instruktionsbladet). Fyra talare användes (2 män; 2 kvinnor) och taldata valdes så att alla yttranden var helt flytande med undantag av FP och ofyllda pauser. För en detaljerad beskrivning, se [Eklund \(2004:187–189\)](#).

Försökspersonerna var 16 friska vuxna med åldersspannet 22–54 (medelålder 40, 3 med en standardavvikelse på 9,5). Alla försökspersoner var högerhänta enligt Edinburgh Handedness Inventory (Oldfield, 1971).

Instruktionerna som gavs till försökspersonerna var att de skulle spela rollen av resebyråagenten och noga lyssna på ”kunden” sade (”jag vill boka en bil den 3 maj”, etc), men att de inte förväntades svara utan enbart tyst skulle lägga uppmärksamhet på vad som sades.

För alla övriga tekniska aspekter (beskrivning av scanner och annan utrustning, experimentdesign, MRI-scanningsmetodik, data-post-processning, etc, hänvisas till Eklund och Ingvar, 2016; länk i slutet av denna artikel).

För en grafisk beskrivning av experimentdesignen, se Bild 1.

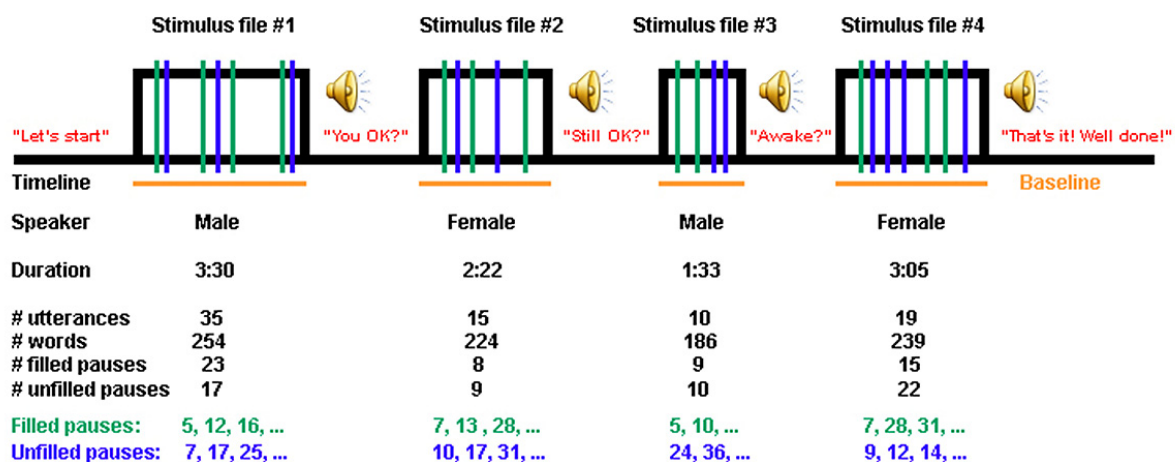


Bild 1: Experimentdesign. Försökspersonerna (fpp) låg i en fMRI-scanner och lyssnade via högtalare på fyra olika ljudfiler som spelades upp en efter en, med en liten verbal kontroll av experimentledaren mellan varje fil att fpp fortfarande var vakna (trots en enorm ljudnivå i en scanner förekommer det att fpp somnar). Stimuli spelades upp med varannan röst manlig, varannan kvinnlig och försökspersonernas hjärnaktivitet registrerades där den orangea linjen anges ”baslinje”-aktivitet (flytande tal), och de gröna siffrorna anger hur långt in i filen (i sekunder) FP förekom, och blå siffror anger detsamma för ofyllda pauser. Eftersom tidsupplösningen inom fMRI inte (som EEG) sker i realtid räcker en precision på sekundnivå. De svarta siffrorna anger durationen för varje fil, hur många yttranden, ord, FP och UP som varje fil innehöll. Genom att kontrastera hjärnaktiviteten för perception mellan olika kategorierna (flytande tal, FP, UP) skapades kontraster vilka utgör resultaten för studien.

Analys och resultat

För analys av data skapades tre kontraster, där flytande tal (FS) utgjorde baslinjen (se Bild 1):

1. FP ökad aktivitet jämfört FS
2. UP ökad aktivitet jämfört FS
3. FP ökad aktivitet jämfört UP

Givet att denna studie är den första (och hittills enda?) i sitt slag tittade vi på hela hjärnan (i stället för att fokusera på så kallade ROI, Regions Of Interest). Resultaten beräknades med en False Discovery Rate med $p < 0,05$ (se [Genovese, Lazar & Nichols, 2002](#)). Tröskeln för klusternivån sattes till tio sammanhängande voxlar.

Vi fann ingen aktivitet i Brodmann Area 22 (BA22), dvs Wernickes Area, vilket förknippas med processning av semantisk information (betydelse).

För **kontrasten FP > FS** fann vi ökad aktivitet i Primära Hörselkortex (PAC) ([Morosan et al., 2001](#); [Rademacher et al. 2001](#)), binauralt, i subkortikala områden (cerebellum, putamen); se Bild 2.

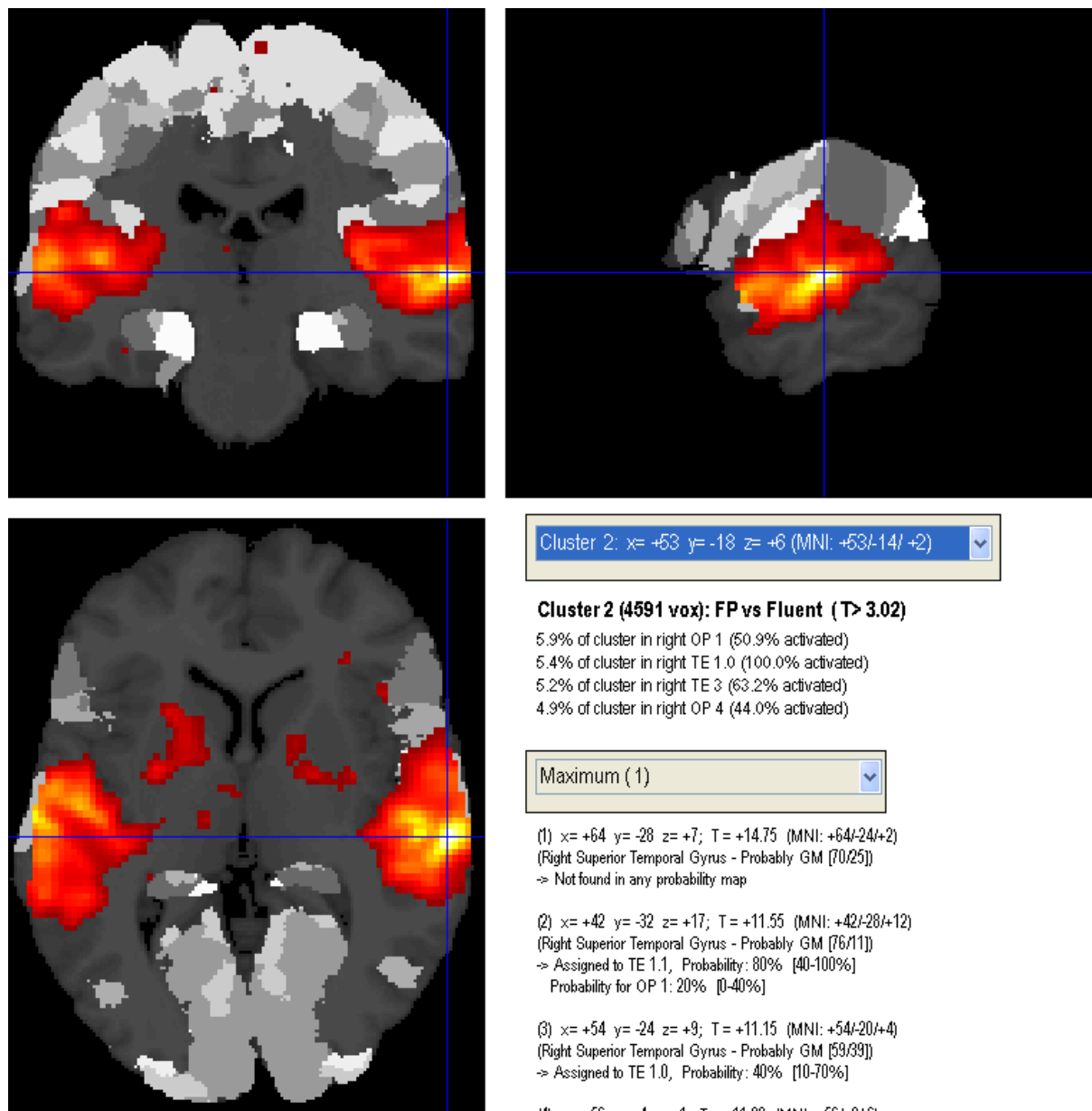


Bild 2: Modulering för kontrasten FP > FS. Aktivering i STG/auditiva kortex.

Intressantare fann vi modulering av Supplementära Motorkortex (SMA), BA6. Typisk aktivering visas i Bild 3 och Bild 4.

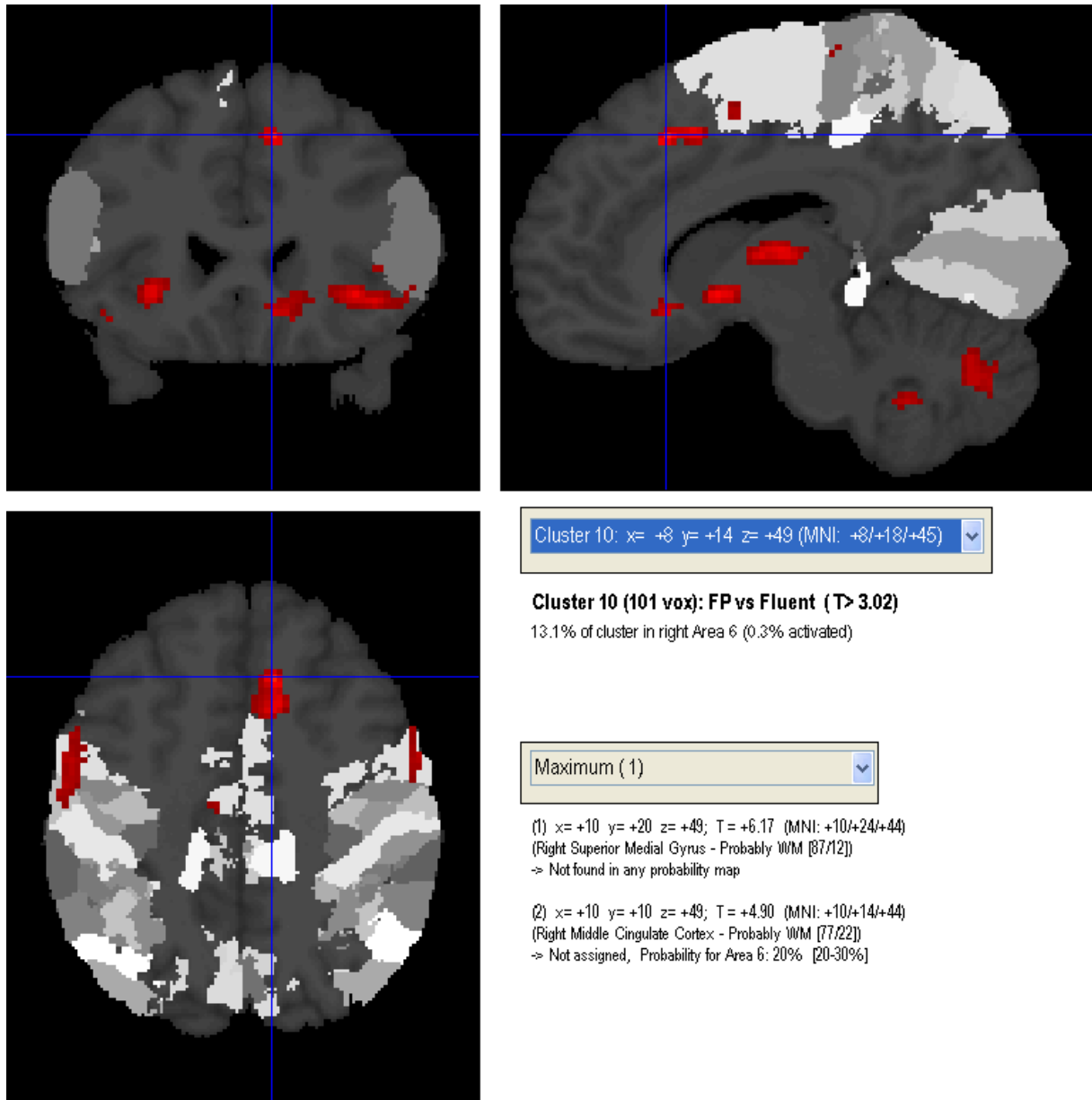


Bild 3: Aktivering för FP > FS. Aktivering i motorkortex (Area 6).

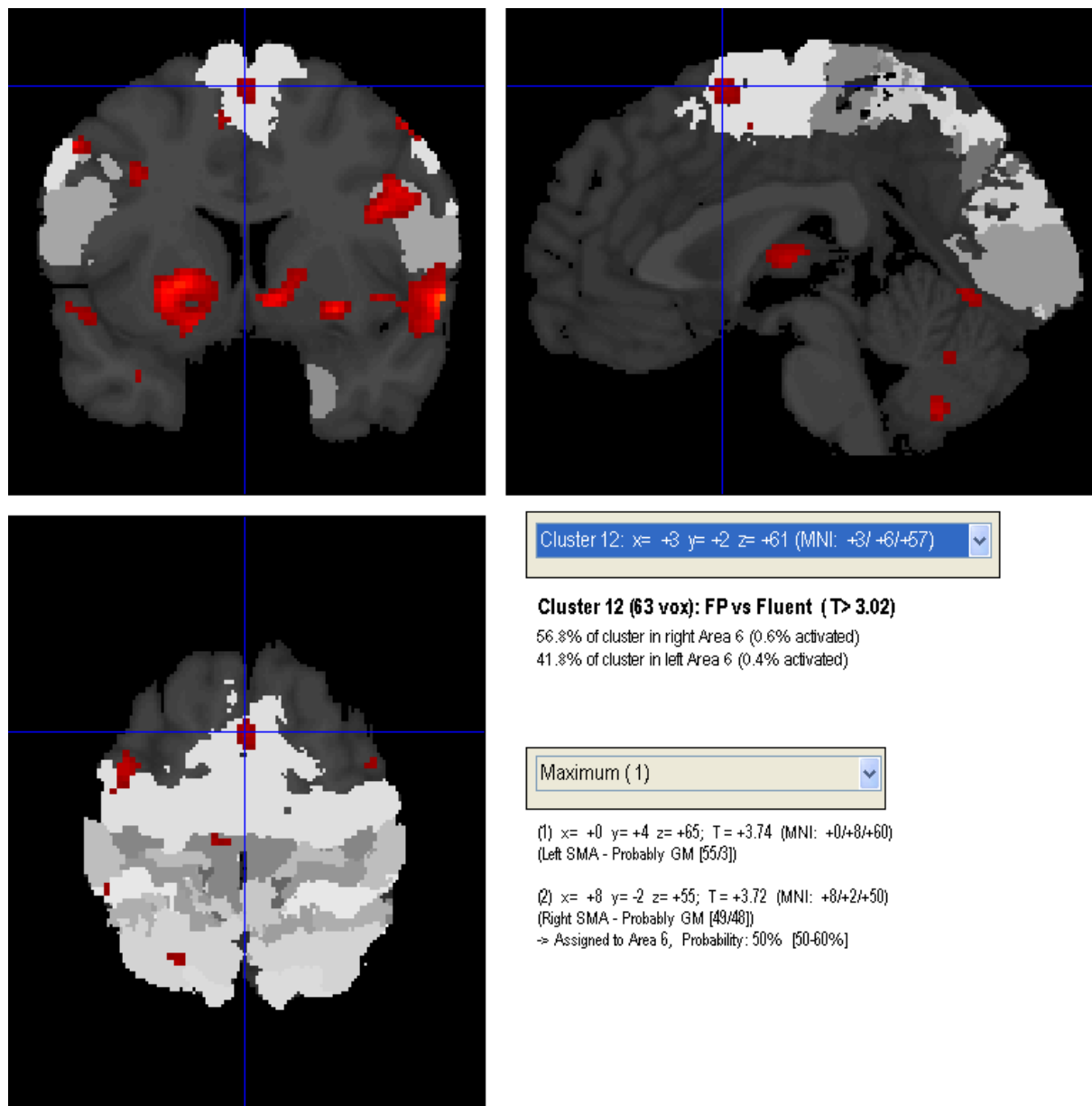


Bild 4: Aktivering för FP > FS. Aktivering i motorkortex (Area 6).

För **kontrasten UP > FS** fann vi ökad aktivering PAC, samt även (bland annat) i Heschl's Gyrus, Rolandiska Operculum. Ingen aktivering i motorkortex observerades. Typisk aktivering visas i Bild 5.

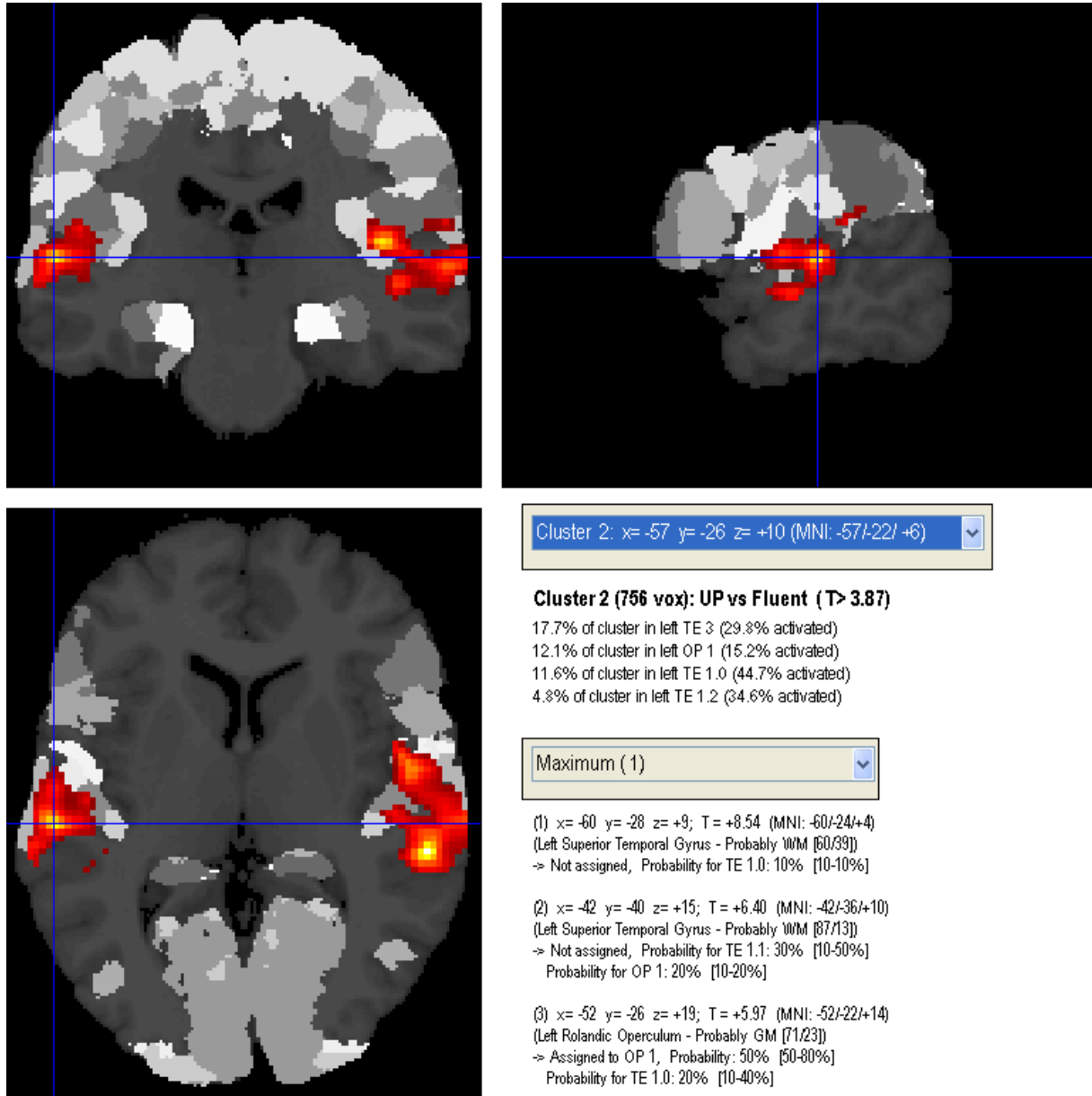


Bild 5: Aktivering för UP > FS. Aktivering i STG/auditiva kortex.

För **kontrasten FP > UP** fanns vi att aktiveringsmönstret till stor liknande FP > FS. Typisk aktivering (endast *ett* exempel) visas i Bild 6.

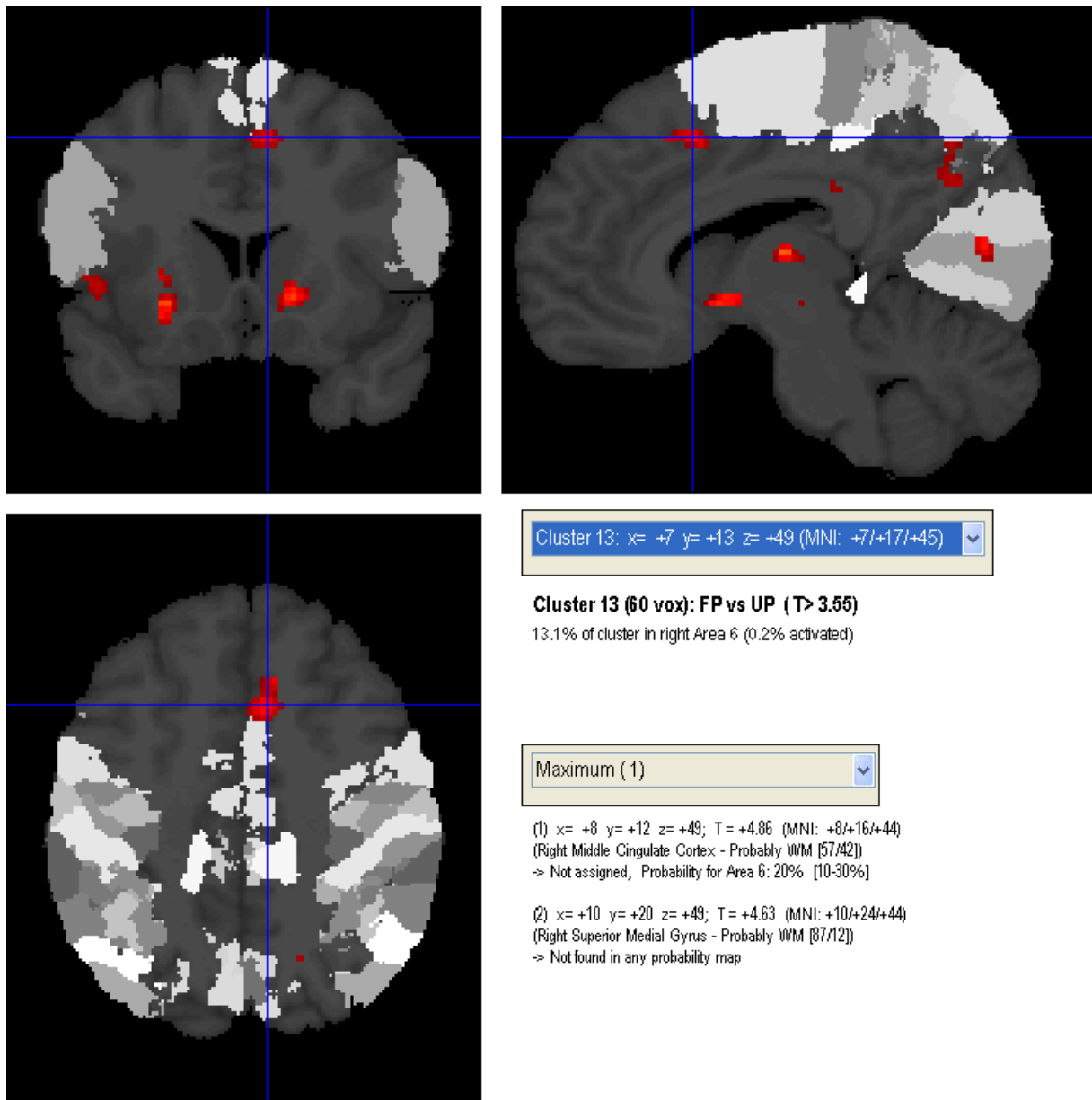


Bild 6: Aktivering för FP > UP. Aktivering i motorkortex (Area 6).

Resultaten visar således att såväl FP som UP påverkar vår auditiva uppmärksamhet (båda aktiverar hörselkortex), men medan UP modulerar områden för syntaxprocessning – se [Eklund och Ingvar \(2016\)](#) för en tabulering av den fullständiga moduleringen av olika områden – så har FP inte denna effekt. FP modulerar i stället motorkortex. Jämfört med FP så verkar FS och UP ha ungefär samma effekt i det att ingen observerad skillnad föreligger i det två kontrasterna FP > FS och FP > UP.

Diskussion

Våra starkaste resultat var *aktiveringen av PAC*, som i hög grad moduleras av såväl UP som FP, dvs lyssnarna verkar ”spetsa öronen” både när de hör UP och när de hör FP. Att ökad uppmärksamhet påverkar perception har tidigare visats (Petkov et al., 2004). Den observerade ökade aktiviteten i PAC kan förklara (åtminstone delvis) de kortade reaktionstider till lingvistiska stimuli som följer FP som rapporterats av Fox Tree (1995, 2001). Eftersom även UP modulerar PAC verkar det emellertid som om även detta ”avbrott” i talsignalen har denna uppmärksamhetsökande effekt, och den hypotes som följer av våra resultat skulle således vara att även UP skulle leda till kortare reaktionstider i experiment av den typ som Fox Tree utförde.

Vad rör vår den *ökade aktiviteten i motorkortex (SMA/BA6)* så finns det flera tänkbara förklaringar. Den mest uppenbara av dessa är att när en lyssnare hör en talare säga ”eh” så förbereder sig lyssnaren för att själv tala. Det är känt sedan minst Tuthill (1944) att SMA är involverat i talprocessning, och senare studier har visat att såväl SMA och pre-SMA spelar en roll vid såväl *talproduktion* (Goldberg, 1985; Alario et al., 2006) som *talperception* (Iacoboni, 2008; Wilson et al. 2004). Det kan emellertid tänkas att SMA aktiverades som en del av talandning, skiljt från vanlig baslinjeandning (Murphy et al. 1997).

Våra resultat har direkta *implikationer för två populära FP-hypoteser*. Vad rör den redan 1959 föreslagna ”floor-holding”-hypotesen (Maclay & Osgood, 1959), som går ut på att talare säger ”eh” för att förhindra lyssnaren att ta över konversation. Detta kan mycket väl stämma, men våra resultat tyder på att produktion av FP i så fall verkar kontraproduktivt: en lyssnare som hör ett ”eh” uppvisar i stället ett beteende, på neural nivå, som tyder på att talproduktionen snarare *aktiveras* än förhindras.

Var rör den alternativa ”help-me-out”-hypotesen (Clark & Wilkes-Gibbs, 1986) så är nyheterna bättre. Om FP aktiverar SMA hos lyssnaren så kommer en interlokutör som hör ett ”eh” antagligen att reagera snabbare på den eftersökta hjälpen.

Sammanfattning

Den aktuella studien är intressant vi tre huvudsakliga skäl:

1. Vi använder *fMRI för att studera disfluensperception*; tidigare studier har huvudsakligen använt sig av EEG med därmed associerat fokus på temporala aspekter av talperception.
2. Vi undersökte den perceptuella modulering orsakad av FP “per se”, inte på (som i tidigare forskning) dessas effekt på följande lingvistiska enheter (ord, fraser) eller allmän kognitiv processning.
3. Till skillnad från tidigare studier där auditiva stimuli ofta har utgjorts av scriptat laboratorietal använde vi ekologiskt valida data, dvs “spontal” insamlat under ett Wizard-of-Oz-paradigm.

Våra resultat är, generellt sett, svåra att jämföra med andra studier, främst på grund av att vi inte känner till att någon (ens avläggset) liknande studie har utförts. Det gör att såväl våra resultat som spekulationer får förbli ”förslag” i väntan på bekräftelser från framtida studier. Våra resultat visar dock att såväl UP som FP får lyssnare att ”spetsa öronen” samt att FP dessutom ”kick-startar” talproduktionen hos lyssnaren, och att FP återigen har visat sig ”leva sitt eget liv”, som George Mahl upptäckte redan i mitten av 1950-talet.

Tack

Tack till Karolinska Institutets etikkommitté, som godkände studien and 4 april 2007. Data gjordes tillgängliga för avslutande av studien via avtal mellan TeliaSonera, Karolinska Institutet och författaren (RE). Djupa tack till min handledare Martin Ingvar, en inspirationsspruta utan dess like. Oerhörda tack till Peter Fransson för hjälp med allt som har med experimentdesign och tekniska aspekter att göra. Tack till mina två ”scannings-kamrater” Örjan de Manzano (né Blom) och Katarina Gospic. Vi hade urkul! Vidare tack att en mängd andra människor ”i korridoren/fikarummet”, inkluderande (men inte uttömmande) Karl-Magnus Peterson, Urban Hansson, Anders Nordell, Karin Jensen, Julia Uddén, Christian Forkstam, Yords Österman och Mimmi Wernman.

Referenser

- Ahlsén, Elisabeth & Robert Eklund. 2012. Språk, hjärnan och kognition. I: Jens Allwood & Mikael Jensen (eds.), *Kognitionsvetenskap*. Lund: Sweden: Studentlitteratur, kapitel 6.4, pp. 437–552.
- Alario, Franz.-X. Hanna Chainay, Stéphane Lehericy & Laurent Cohen. 2006. The role of the supplementary motor area (SMA) in word production. *Brain Research*, vol. 1076, pp. 129–143.
- Arnold, Jennifer E., Maria Fagnano & Michael K. Tanenhaus. 2003. Disfluencies Signal Theree, Um, New Information. *Journal of Psycholinguistic Research*, Vol. 32, No. 1, pp. 25–36.
- Ferreira, Fernanda & Karl G.D. Bailey. 2004. Disfluencies and human language comprehension. *TRENDS in Cognitive Sciences*, vol. 8, no. 5), pp. 231–237.
- Baumeister, Roy F. 1984. Choking Under Pressure: Self-Consciousness and Paradoxical Effects of Incentives on Skillful Performance. *Journal of Personality and Social Psychology*, vol. 46, no. 3, pp. 610–620.
- Beattie, George W. & P.J. Barnard. 1979. The temporal structure of natural telephone conversations (directory enquiry calls). *Linguistics*, 17:213–229.
- Bell, Linda, Robert Eklund & Joakim Gustafson. 2000. A Comparison of Disfluency Distribution in a Unimodal and a Multimodal Human–Machine Interface *Proceedings of ICSLP '00*, Beijing, 16–20 October 2000, vol. 3, pp. 626–629.
- Beyerstein, Barry L. 1999: Whence Cometh the Myth that We Only Use 10% of our Brains? IO: Sergio Della Sala (ed): *Mind Myths. Exploring Popular Assumptions About The Mind and Brain*. Chicester: John Wiley and Sons, ch. 1m pp. 3–24.
- Boomer, Donald S. 1965. Hesitation and grammatical encoding. *Language and Speech*, vol. 8, pp. 148–158.
- Bortfeld, Heather, Silvia D. Leon, Jonathan E. Bloom, Michael F. Schober & Susan E. Brennan. 2001. Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech*, vol. 44, no. 2, pp. 123–147.
- Bortfeld, Heather, Silvia D. Leon, Jonathan E. Bloom, Michael F. Schober & Susan E. Brennan. 1999. Which speakers are most disfluent in conversation, and when? *Proceedings of Disfluency in Spontaneous Speech Workshop*, 1 July 1999, Berkeley, California, USA, pp. 7–10.
- Brennan, Susan E. & Michael F. Schober. 2001. How Listeners Compensate for Disfluencies in Spontaneous Speech. *Journal of Memory and Language*, vol. 44, pp. 274–296.
- Dale, Paulette. 1977. Factors Related to Dysfluent Speech in Bilingual Cuban-American Adolescents. *Journal of Fluency Disorders*, vol. 2, pp. 311–314.

- Christenfeld, Nicholas & Beth Creager. 1996. Anxiety, Alcohol, Aphasia, and Ums. *Journal of Personality and Social Psychology*, vol. 70, no. 3, pp. 451–460.
- Christenfeld, Nicholas. 1995. Does It Hurt To Say Um? *Journal of Nonverbal Behavior*, vol. 19, no. 3, pp. 171–186.
- Clark, Herbert H. & Deanna Wilkes-Gibbs. Referring as a collaborative process. 1986. *Cognition*, vol. 22, pp. 1–39.
- Clark, Herbert H. & Jean E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speech. *Cognition*, vol. 84, pp. 73–111.
- Cook, Mark. 1971. The incidence of filled pauses in relation to part of speech. *Language and Speech*, vol. 14, part 2, pp. 135–139.
- Corley, Martin, L.J. MacGregor, D.I. Donaldson. 2007. It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, vol. 105, pp. 658–668.d
- Duffy, Robert J., Martin F. Hunt Jr. & Thomas G. Giolas. 1975. Effects of Four Types of Disfluency on Listener Reactions. *Folia Phoniatrica*, vol. 27, no. 2, pp. 106–115.
- Dündar, Sefa & Nazan Gündüz. 2016. Misconceptions Regarding the Brain: The Neuromyths of Preservice Teachers. *Mind, Brain and Education*, vol. 10, no. 4, pp. 212–232.
- Eklund, Robert & Elizabeth Shriberg. 1998. Crosslinguistic Disfluency Modelling: A Comparative Analysis of Swedish and American English Human–Human and Human–Machine Dialogues. *Proceedings of ICSLP 98*, Sydney, 30 November–5 December 1998. Paper 805, vol. 6, pp. 2631–2634.
- Eklund, Robert. 2004. *Disfluency in Swedish human–human and human–machine travel booking dialogues*. PhD thesis, Linköping Studies in Science and Technology, Dissertation No. 882, Department of Computer and Information Science, Linköping University, Sweden, ISBN 91-7373-966-9, ISSN 0345-7524.
- Eklund, Robert. 2010. The Effect of Directed and Open Disambiguation Prompts in Authentic Call Center Data on the Frequency and Distribution of Filled Pauses and Possible Implications for Filled Pause Hypotheses and Data Collection Methodology. *Proceedings of DiSS-LPSS Joint Workshop 2010, The 5th Workshop on Disfluency in Spontaneous Speech and The 2nd International Symposium on Linguistic Patterns in Spontaneous Speech*. University of Tokyo, Japan, September 25–26 September 2010, Tokyo, Japan, pp. 23–26.
- Eklund, Robert & Mats Wirén. 2010. Effects of open and directed prompts on filled pauses and utterance production. *Proceedings of Fonetik 2010*, Lund University, 2–4 June 2010, pp. 23–28.
- Eklund, Robert & Martin Ingvar. 2016. Supplementary Motor Area Activation in Disfluency Perception. An fMRI Study of Listener Neural Responses to Spontaneously Produced Unfilled and Filled Pauses In: *Proceedings of Interspeech 2016*, 9–12 September 2016, San Francisco, USA, DOI: <http://dx.doi.org/10.21437/Interspeech.2016-1082>, pp. 1378–1381.
- Fox Tree, Jean E. 2001. Listeners' uses of *um* and *uh* in speech comprehension. *Memory and Cognition*, vol. 29, no. 2, pp. 320–236.
- Fox Tree, Jean E. 1995. The Effects of False Starts and Repetitions on the Processing of Subsequent Words in Spontaneous Speech. *Journal of Memory and Language*, vol. 34, pp. 709–728.
- Goldberg, G. 1985. Supplementary motor area structure and function: Review and hypotheses. *The Behavioral and Brain Sciences*, vol. 8, pp. 567–616.
- Genovese, Christopher R. Nicole A. Lazar & Thomas Nichols. Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate, *NeuroImage*, vol. 15, pp. 870–878, 2002
- Goldman-Eisler, Frieda. 1957. Speech Production and Language Statistics. *Nature*, vol. 28, December 1957, p. 1497.
- Goldman-Eisler, Frieda. 1955. Speech-breathing activity—A measure of tension and affect during interviews. *British Journal of Psychology*, vol. 46, pp. 53–63.
- Goldman-Eisler, Frieda. 1954a. A study of individual differences and of interaction in the behaviour of some aspects of language in interviews. *Journal of Mental Science*, vol. 100, pp. 177–197.
- Goldman-Eisler, Frieda. 1954b. On the variability of the speech of talking and on its relation to the length of utterances in conversations. *British Journal of Psychology*, vol. 45, pp. 94–107.
- Hegedüs, Lajos. 1953. On the problem of the pauses of speech. *Acta Linguistica Academiae Scientiarum Hungaricae*, vol. 3, pp. 1–34.
- Hindle, Donald. 1983. Deterministic Parsing of Syntactic Non-fluencies. *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, 15–17 June 1983, Cambridge, Massachusetts, USA, pp. 123–128.
- Iacoboni, Marco. 2008. The role of premotor cortex in speech perception: Evidence from fMRI and rTMS. *Journal of Physiology – Paris*, vol. 102, pp. 31–34.
- Jain, Varsha, Michael C. Langham & Felix W. Wehrli. 2010. MRI estimation of global brain consumption rate. *Journal of Cerebral Blood Flow & Metabolism*, vol. 30, no. 9, pp. 1598–1607.

- Jackowski, Christian, Stephan Bolliger & Michael J. Thali. 2008. Common and Unexpected Findings in Mummies from Ancient Egypt and South America as Revealed by CT. *RadioGraphics*, vol.28, pp. 1477–1492.
- James, William. 1890 (1950). *The Principles of Psychology*. New York: Dover Publications. Facsimile edition of the original Henry Holt (1890) edition.
- Johnson, Wendell, Spencer F. Brown, James F. Curtis, Clarence W. Edney & Jacqueline Keaster. 1948. Stuttering. In: Wendell Johnson, Spencer F. Brown, James F. Curtis, Clarence W. Edney & Jacqueline Keaster (eds.), *Speech Handicapped School Children*. New York: Harper & Brothers Publishers, ch 5, pp. 179–257.
- Johnson, Wendell (ed.). 1955. *Stuttering in Children and Adults. Thirty Years of Research at the University of Iowa*. Minneapolis: University of Iowa Press.
- Johnson, Wendell and Associates. 1959. *The Onset of Stuttering: Research Findings and Implications*. Minneapolis: University of Minnesota Press.
- Johnson, Wendell. 1961. Measurement of Oral Reading and Speaking Rate and Disfluency of Adult Male and Female Stutterers and Nonstutterers. *Journal of Speech and Hearing Disorders Monograph Supplement Number 7*, pp. 1–20.
- Kasl, Stanislav V. & George F. Mahl. 1958. Experimentally induced anxiety and speech disturbances. *American Psychologist*, vol. 13, p. 349.
- Kasl, Stanislav V. & George F. Mahl. 1965. The relationship of disturbances and hesitations in spontaneous speech to anxiety. *Journal of Personality and Social Psychology*, vol. 1, pp. 425–433.
- Kasl, Stanislav V. & George F. Mahl. 1987. Speech Disturbances and Experimentally Induced Anxiety. In: George F. Mahl (ed.), *Explorations in Nonverbal and Vocal Behavior*. Hillsdale, New Jersey: Lawrence Erlbaum, ch. 12, pp. 203–213.
- Lalljee, Mansur [G.] & Mark Cook. 1973. Uncertainty in first encounters. *Journal of Personality and Social Psychology*, vol. 26, no. 1, pp. 137–141.
- Lalljee, Mansur G. & Mark Cook. 1969. An experimental investigation of the function of filled pauses in speech. *Language and Speech*, vol. 12, pt. 1, pp. 24–28.
- Lickley, R[obin]. J. & E[llen]. G[urman]. Bard. 1996. On not Recognizing Disfluencies in Dialogue. *Proceedings of the International Conference on Spoken Language Processing (ICSLP) '96*, 3–6 October 1996, Philadelphia, Pennsylvania, USA, vol. 3, pp. 1876–1879.
- Livant, William Paul. 1963. Antagonistic functions of verbal pauses: filled and unfilled pauses in the solution of additions. *Language and Speech*, vol. 6, part 1, pp. 1–4.
- Lounsbury, Floyd G. 1954. Transitional Probability, Linguistic Structure, and Systems of Habit-family Hierarchies. In: Charles E. Osgood & Thomas A. Sebeok (eds.), *Psycholinguistics. A Survey of Theory and Research Problems*. Baltimore: Waverly Press, pp. 93–101.
- Lutz, Konnie C. & A. R. Mallard. 1986. Disfluencies and rate of speech in young adult nonstutterers. *Journal of Disfluency Disorders*, 11:307–316.
- Maclay, Howard & Charles E. Osgood. 1959. Hesitation Phenomena in Spontaneous English Speech. *Word*, vol. 5, pp. 19–44.
- Mahl, George F. 1956. Disturbances and silences in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology*, vol. 53, pp. 1–15.
- Mahl, George F. 1958. On the use of “ah” in spontaneous speech: Quantitative, developmental, characterological, situational, and linguistic aspects. *American Psychologist*, vol. 13, p. 349.
- Mahl, George F. (ed.). 1987a. *Explorations in nonverbal and vocal Behavior*. Hillsdale, New Jersey: Lawrence Erlbaum.
- Mahl, George F. 1987b. Everyday Disturbances in Speech. In: Robert L. Russell (ed.), *Language in Psychotherapy: Strategies of Discovery*. New York and London: Plenum Press, ch. 6, pp. 213–269.
- McCroskey, James C. & R. Samuel Mehrley. 1969. The effects of disorganization and nonfluency on attitude and source credibility. *Speech Monographs*, vol. 36, pp. 13–21.
- Moniz, Helena, Ana Isabel Mata & M. Céu Viana. 2007. On Filled-Pauses and Prolongations in European Portuguese. *Proceeding of Interspeech 2007*, 27–31 August 2007, Antwerp, 2645–2648.
- Morosan, P., J. Rademacher, A. Schleicher, K. Amunts, T. Schormann, K. Zilles. 2001. Human Primary Auditory Cortex: Cytoarchitectonic Subdivisions and Mapping into a Spatial Reference System. *NeuroImage*, vol. 13, pp. 684–701.
- Murphy, K., D. R. Corfield, A. Guz, G. R. Fink, R. J. S. Wise, K. Harrison, L. Adams. 1997. Cerebral areas associated with motor control of speech in humans. *Journal of Applied Physiology* vol. 83, no. 5, pp. 1438–1447.
- O'Connell, Daniel C. & Sabine Kowal. 2005. Uh and Um Revisited: Are They Interjections for Signaling Delay? *Journal of Psycholinguistic Research*, vol. 34, no. 6, pp. 575–576.

- Oldfield, R. C. 1971. The Assessment and Analysis of Handedness: The Edinburgh Inventory. *Neuropsychologia*, vol. 9, pp. 97–113.
- Oviatt, Sharon. 1995. Predicting spoken disfluencies during human–computer interaction. *Computer Speech and Language*, vol. 9, pp. 19–35.
- Petkov, Christopher I., Xiaoiang Kang, Kimmo Alho, Olivier Bertrand, E. William Yund & David L. Woods. 2004. Attentional modulation of human auditory cortex. *Nature Neuroscience*, vol. 7, no. 6, pp. 658–663.
- Pfeifer, L.M. & T. Bickmore. 2009. Should Agents Speak, Like, um, Humans? The Use of Conversational Fillers by Virtual Agents. I: Z. Ruttkay et al. (eds.): *Intelligent Virtual Agents. Proc. 9th International Conference on Intelligent Virtual Agents (IVA)*, 14–16 September 2009, Amsterdam, 460–466.
- Rademacher, J., P. Morosan, T. Schormann, A. Schleicher, C. Werner, H.-J. Freund, K. Zilles. 2001. Probabilistic Mapping and Volume Measurement of Human Primary Auditory Cortex. *NeuroImage*, vol. 13, pp. 669–683.
- Ratner, Nan Bernstein & Catherine Costa Sih. 1987. Effects of gradual increases in sentence length and complexity on children’s dysfluency. *Journal of Speech and Hearing Disorders*, vol. 52, pp. 278–287.
- Schachter, Stanley, Nicholas Christenfeld, Bernard Ravina & Frances Bilous. 1991. Speech Disfluency and the Structure of Knowledge. *Journal of Personality and Social Psychology*, vol. 60, no. 3, pp. 362–367.
- Shapiro, Arnold I. & Barbara A. DeCicco. 1982. The relationship between normal dysfluency and stuttering: An old question revisited. *Journal of Fluency Disorders*, vol. 7, pp. 109–121.
- Siegel, Gerald M., Joanne Lenske & Patricia Broen. 1969. Suppression of normal speech disfluencies through response costs. *Journal of Applied Behavior Analysis*, vol. 2, p. 265–276.
- Shriberg, Elizabeth Ellen. 1994. *Preliminaries to a Theory of Speech Disfluencies*. PhD thesis, University of California, Berkeley.
- Tuthill, Curtis E. 1940. A quantitative study of extensional meaning with special reference to stuttering. *Journal of Speech Disorders*, vol. 5, pp. 189–191.
- Ward, Jamie. 2006. *The Student’s Guide to Cognitive Neuroscience*. Hove and New York: Psychology Press.
- Westby, Carol E. 1974. Language performance of stuttering and nonstuttering children. *Journal of Communications Disorders*, vol. 12, pp. 133–145.
- Wanjek, C. 2002. 10 percent misconception, 90 percent misdirection: The brain at work. I: C. Wanjek (ed.), *Bad medicine: Misconceptions and misuses revealed, from Distance Healing to Vitamin O*. New York, NY: Wiley., ch.1, pp. 17–22.
- Wilson, Stephen M., Ayşe Pinar Saygın, Martin I. Sereno & Marco Iacoboni. 2004. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, vol. 7, no. 7, pp. 701–702.
- Wingate, Marcel E. 1984b. Fluency, disfluency, dysfluency, and stuttering. *Journal of Fluency Disorders*, vol. 17, pp. 163–168.

Länkar

Proceedings från alla hittills organiserade DiSS-workshops kan laddas ned via denna länk:

<https://roberteklund.info/conferences/diss2017/Proceedings.htm>

Eklund & Ingvar (2016) kan laddas ned via denna länk:

https://www.ida.liu.se/~robek28/pdf/Eklund_Ingvar_2016_Disfluency_fMRI_FilledPauses.pdf

Radiointervjuer om fyllda pausers funktion i radiotal:

NRK *Språkteigen*, 22 november 2009:

https://www.ida.liu.se/~robek28/audio/disfluency-NorwegianRadio_Sprakteigen_2009-11-22.mp3

SR *Språket*, 1 June 2010:

https://www.ida.liu.se/~robek28/audio/disfluency-SwedishRadio_Spraket_2010-06-01_FilledPauses.mp3

”Crash course” i fMRI:

<http://www.fmri4newbies.com/>